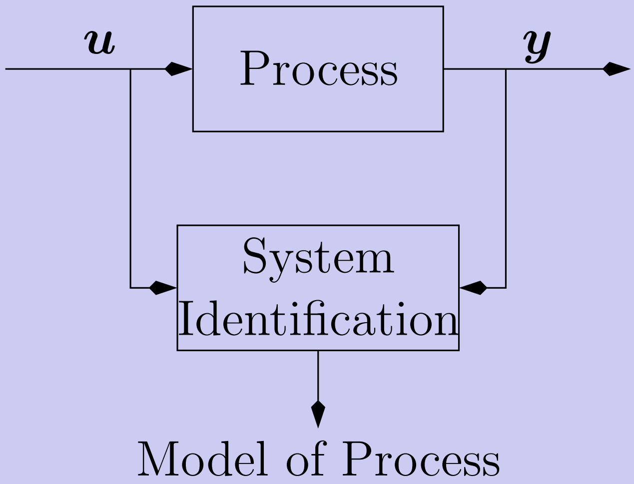
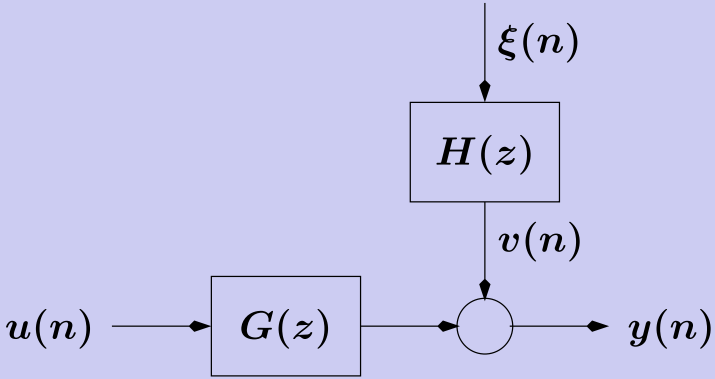
1. System Identification Problem



- $u(n)$ is control input
- $\xi(n)$ is white noise
- Would like to determine of transfer functions $G(z)$ and $H(z)$ from measured data

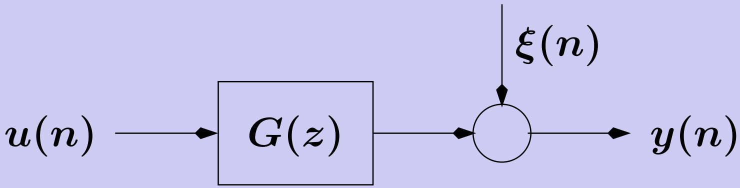
2. System Identification Problem: A Model



- $u(n)$ is control input, $\xi(n)$ is white noise
- Would like to determine plant transfer function $G(z)$ and disturbance transfer function $H(z)$ from data
- Difficult to converge to solution - propose a two step approach

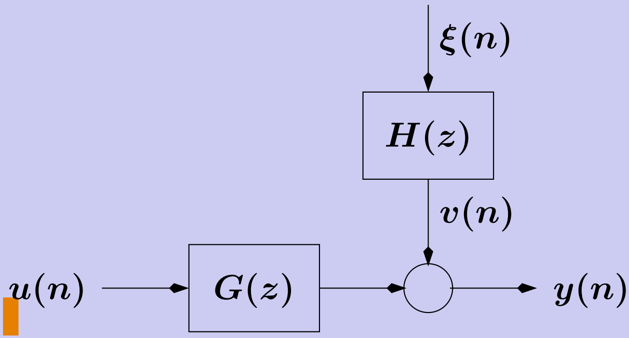
3. System Identification Problem in Detail

- First find G only, assuming noise to be white:



- Check if $y - Gu$ is white. If it is white, done.

  If not, let $y - Gu = v$



- Determine $H(z)$ between v and white noise $\xi(n)$

4. Stochastic Processes

- A stochastic process is a statistical phenomenon that evolves in time according to probabilistic laws.
- A realization is a sample of the many possibilities that a process can take (population).
- A time-series is a set of values sampled from the process sequentially.
- A time-series is a particular realization of the process.

5. Mean, ACF, CCF of Stationary, Ergodic Processes

Mean

$$m_u = \frac{1}{2N+1} \sum_{n=-N}^{N} u(n)$$

Auto Covariance Function (ACF):

$$r_{uu}(l) = \frac{1}{2N+1} \sum_{k=-N}^{N} (u(k) - m_u)(u(k-l) - m_u)$$

Cross Covariance Function (CCF)

$$r_{uy}(l) = \frac{1}{2N+1} \sum_{k=-N}^{N} (u(k) - m_u)(y(k-l) - m_y)$$

6. White Noise

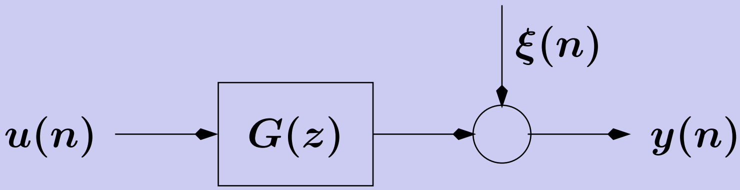
- The discrete-time white-noise sequence $\{\xi(k)\}$ is a set of independent, identically distributed (i.i.d.) values belonging to a stationary stochastic process.
- The mean of white noise is zero.
- White noise has infinite energy $\Rightarrow$ no Fourier Transform.
- The ACF of a white-noise sequence is given by:

$$r_{\xi\xi}(k) = \sigma_\xi^2 \delta(k) = \begin{cases} \sigma_\xi^2 & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

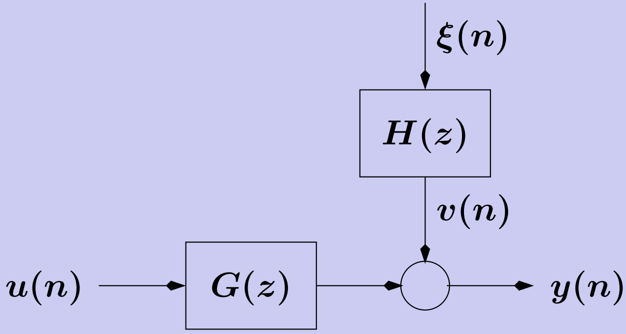
- The Z-transform of ACF of white noise is σ_ξ^2.
- Fourier Transform of ACF (power spectrum) of white-noise is constant and given by, $\Phi_{\xi\xi}(\omega) = \sigma_\xi^2, \quad \forall \omega$.

7. System Identification Problem in Detail

- First find G only, assuming noise to be white:



- Check if $y - Gu$ is white. If it is white, done.



If not, let $y - Gu = v$

- Determine $H(z)$ between v and white noise $\xi(n)$

8. Mixed Notation

Recall the model of our system:

$$y(n) = g(n) * u(n) + \xi(n)$$

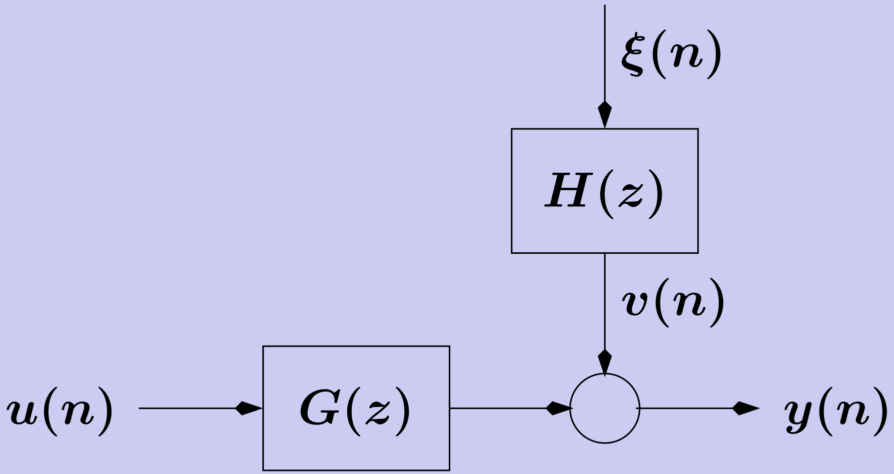
- We have difficulty in taking Z-transform of this equation.
- The Z-transform of $\xi(n)$ may not exist.
- It is inconvenient to carry the convolution operation.
- In view of this, a mixed notation is used

$$y(n) = G(z)u(n) + \xi(n)$$

$G(z)$ denotes the Z-transform of $g(n)$.

$$\begin{aligned} y(n) &= g(n) * u(n) + h(n) * \xi(n) \\ &= G(z)u(n) + H(z)\xi(n) \end{aligned}$$

9. Modelling H with ARMA



ARMA: Auto Regressive Moving Average.

Required to model noise process.

$$
\begin{aligned}
v(n) + a_1 v(n-1) + \cdots + a_p v(n-p) = \\
\xi(n) + c_1 \xi(n-1) + \cdots + c_q \xi(n-q)
\end{aligned}
$$

10. ARMA, AR Processes

Recall ARMA

$$v(n) + a_1 v(n-1) + \cdots + a_p v(n-p) = \xi(n) + c_1 \xi(n-1) + \cdots + c_q \xi(n-q)$$

If $q = 0$, obtain an AR(p) process:

$$v(n) + \cdots + a_p v(n-p) = \xi(n) = \left(1 + \sum_{k=1}^{p} a_k z^{-k}\right) v(n$$

or, equivalently,

$$v(n) = \frac{1}{\left(1 + \sum_{k=1}^{p} a_k z^{-k}\right)} \xi(n) = \frac{1}{A(z)} \xi(n)$$

11. ARMA, MA Processes

Recall ARMA

$$
\begin{aligned}
v(n) + a_1 v(n-1) + \cdots + a_p v(n-p) = \\
\xi(n) + c_1 \xi(n-1) + \cdots + c_q \xi(n-q)
\end{aligned}
$$

When $p = 0$, arrive at MA(q) process:

$$
\begin{aligned}
v(n) &= \xi(n) + c_1 \xi(n-1) + \cdots + c_q \xi(n-q) \\
&= \left(1 + \sum_{k=1}^{q} c_k z^{-k}\right) \xi(n) = C(z)\xi(n)
\end{aligned}
$$

12. ARMA Processes

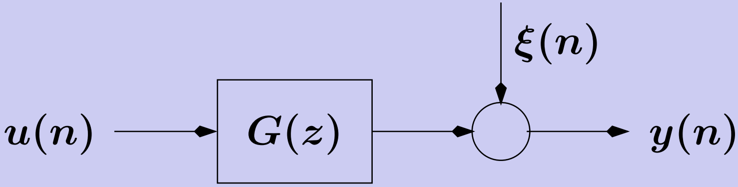
$$v(n) + a_1 v(n-1) + \cdots + a_p v(n-p) = \xi(n) + c_1 \xi(n-1) + \cdots + c_q \xi(n-q)$$

ARMA contains both AR and MA components. Using the procedure explained above,

$$v(n) = \frac{C(z)}{A(z)}\xi(n) = \frac{1 + \sum_{n=1}^{q} c_n z^{-n}}{1 + \sum_{n=1}^{p} a_n z^{-n}}\xi(n)$$

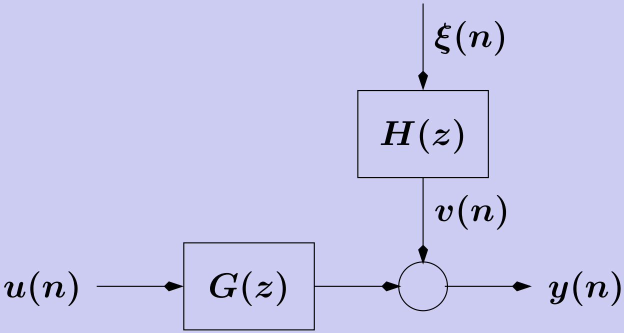
13. Consistent Estimation of G and H

- Assuming white, find G only:



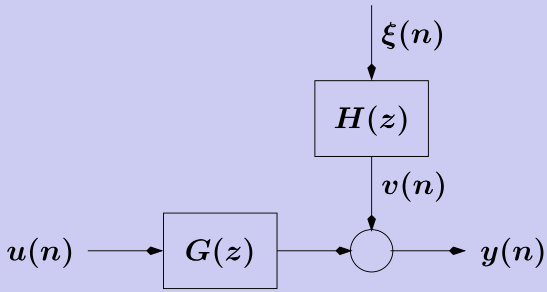
- Check if $y - Gu$ is white. If it is white, done.

If not, let $y - Gu = v$



- Determine $H(z)$ between v and white noise $\xi(n)$
- Use these G and H as the initial guesses in a simultaneous determination approach - one step ahead prediction method

14. One Step Ahead Prediction Error Model



$$y(n) = G(z)u(n) + v(n)$$

Because v is random, need to replace with estimates. The one step ahead prediction error model is given by

$$\begin{aligned}\hat{y}(n|n-1) = H^{-1}(z)G(z)u(n) \\ + [1 - H^{-1}(z)]y(n)\end{aligned}$$

15. General Model

$$A(z)y(n) = \frac{B(z)}{F(z)}u(n) + \frac{C(z)}{D(z)}\xi(n)$$

where $A(z)$, $B(z)$, $C(z)$, $D(z)$ and $F(z)$ are polynomials in z^{-1} defined as

$$\begin{aligned}
A(z) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \cdots + a_{\mathrm{d}A} z^{-\mathrm{d}A} \\
B(z) &= \qquad b_1 z^{-1} + b_2 z^{-2} + \cdots + a_{\mathrm{d}B} z^{-\mathrm{d}B} \\
C(z) &= 1 + c_1 z^{-1} + c_2 z^{-2} + \cdots + a_{\mathrm{d}C} z^{-\mathrm{d}C} \\
D(z) &= 1 + d_1 z^{-1} + d_2 z^{-2} + \cdots + a_{\mathrm{d}D} z^{-\mathrm{d}D} \\
F(z) &= 1 + f_1 z^{-1} + f_2 z^{-2} + \cdots + a_{\mathrm{d}F} z^{-\mathrm{d}F}
\end{aligned}$$

Constant term of B is zero. Transfer function is strictly rational. At least one sample delay before input affects output.

16. One Step Ahead PEM - Examples

Model:

$$y(k) = G(z)u(k) + H(z)\xi(k)$$

Prediction model:

$$\hat{y}(k|k-1) = H^{-1}(z)G(z)u(k) + [1 - H^{-1}(z)]y(k)$$

FIR model:

$$y(k) = B(z)u(k) + \xi(k)$$

It has the following values:

$$G(z) = B(z), \; H(z) = 1$$

Substituting, prediction model for FIR:

$$\hat{y}(k|k-1) = B(z)u(k)$$

17. FIR Model as a Regression Equation

$$y(n) = g(n) * u(n), \quad y(k) = \sum_{l=0}^{N} g(l)u(k-l) + \xi(k)$$

Writing for $y(k)$, $y(k-1)$, $\ldots$ and stacking them,

$$\begin{bmatrix} y(k) \\ y(k-1) \\ \vdots \end{bmatrix} = \begin{bmatrix} u(k) & \cdots & u(k-N) \\ u(k-1) & \cdots & u(k-N-1) \\ \vdots & & \end{bmatrix} \begin{bmatrix} g(0) \\ g(1) \\ \vdots \\ g(N) \end{bmatrix}$$

$$+ \begin{bmatrix} \xi(k) \\ \xi(k-1) \\ \vdots \end{bmatrix}$$

$Z(k) \stackrel{\triangle}{=} \Phi(k)\theta + \Xi(k)$, θ has unknowns

18. One Step Ahead PEM - Examples

Recall Prediction model:

$$\hat{y}(k|k-1) = H^{-1}(z)G(z)u(k) + [1 - H^{-1}(z)]y(k)$$

ARX model:

$$A(z)y(k) = B(z)u(k) + \xi(k)$$

Obtain,

$$G(z) = \frac{B(z)}{A(z)}, H(z) = \frac{1}{A(z)}$$

Substituting, prediction model for ARX:

$$\begin{aligned}\hat{y}(k|k-1) &= A(z)\frac{B(z)}{A(z)}u(k) + (1 - A(z))y(k) \\ &= B(z)u(k) + (1 - A(z))y(k)\end{aligned}$$

19. Models of Interest

- Finite Impulse Response (FIR) model, which is of the form,

$$y(n) = B(z)u(n) + \xi(n)$$

- Auto Regressive with eXogeneous input (ARX) model, which is of the form,

$$A(z)y(n) = B(z)u(n) + \xi(n)$$

- Auto Regressive Moving Average with eXogeneous (ARMAX) model, which is of the form,

$$A(z)y(n) = B(z)u(n) + C(z)\xi(n)$$

20. Models of Interest - Continued

- Auto Regressive Integrated Moving Average with eXogeneous (ARIMAX) model, which is of the form,

$$A(z)y(n) = B(z)u(n) + \frac{C(z)}{\Delta(z)}\xi(n)$$

where, $\Delta = 1 - z^{-1}$.

- Output Error (OE) model, the general form of which is,

$$y(n) = G(z)u(n) + \xi(n)$$

where, G is a transfer function. FIR is an OE model.

- Box Jenkins (BJ) model, which is of the form,

$$y(n) = G(z)u(n) + H(z)\xi(n)$$

$G(z)$ and $H(z)$ are transfer functions.

21. Least Squares Estimation: Regression Equation

- Least Squares Estimation is a convenient method to determine model parameters from experimental data.
- Let the model that relates the parameters and experimental data be given by

$$\boldsymbol{Z(k) = \Phi(k)\theta + \Xi(k)}.$$

- $\boldsymbol{Z(k)}$ and $\boldsymbol{\Phi(k)}$ consist of measurements and $\boldsymbol{\theta}$ is a vector of parameters to be estimated.
- $\boldsymbol{\Xi(k)}$ can be thought of as a mismatch between the best that the underlying model, characterized by $\boldsymbol{\theta}$, can predict and the actual measurement $\boldsymbol{Z(k)}$.

22. Least Squares Estimation: Regression Equation

$$Z(k) = \Phi(k)\theta + \Xi(k)$$

- $\Xi(k)$ can also be thought of as random measurement noise.
- Known as the regression equation.
- Argument k is required in identification problems that received data on a continuous basis.
- If the problem at hand is to determine a set of parameters θ from one and only set of experimental data, there is no need to include this argument.

23. Solution to Least Squares Problem

Regression equation:

$$\boldsymbol{Z}(\boldsymbol{k}) = \boldsymbol{\Phi}(\boldsymbol{k})\boldsymbol{\theta} + \boldsymbol{\Xi}(\boldsymbol{k})$$

Assume $\boldsymbol{E}$ to be negligible. Model:

$$\hat{\boldsymbol{Z}}(\boldsymbol{k}) = \boldsymbol{\Phi}(\boldsymbol{k})\hat{\boldsymbol{\theta}}(\boldsymbol{k})$$

$\hat{\boldsymbol{\theta}}(\boldsymbol{k})$: estimate. Error:

$$\tilde{\boldsymbol{Z}}(\boldsymbol{k}) \stackrel{\triangle}{=} \boldsymbol{Z}(\boldsymbol{k}) - \hat{\boldsymbol{Z}}(\boldsymbol{k})$$

Want $\tilde{\boldsymbol{Z}}$ to be small

24. Solution to Least Squares Problem

2×2 example:

$$Z(k) = \begin{bmatrix} y(k) \\ y(k-1) \end{bmatrix}, \hat{Z}(k) = \begin{bmatrix} \hat{y}(k) \\ \hat{y}(k-1) \end{bmatrix}$$

$$\tilde{Z}(k) = \begin{bmatrix} \tilde{z}(k) = y(k) - \hat{y}(k) \\ \tilde{z}(k-1) = y(k-1) - \hat{y}(k-1) \end{bmatrix}$$

Form an objective function to minimize:

$$\begin{aligned} \tilde{Z}^T(k) W(k) \tilde{Z}(k) &= \begin{bmatrix} \tilde{z}(k) & \tilde{z}(k-1) \end{bmatrix} \\ &\quad \begin{bmatrix} w(k) & 0 \\ 0 & w(k-1) \end{bmatrix} \begin{bmatrix} \tilde{z}(k) \\ \tilde{z}(k-1) \end{bmatrix} \\ &= \begin{bmatrix} \tilde{z}(k) & \tilde{z}(k-1) \end{bmatrix} \begin{bmatrix} w(k)\tilde{z}(k) \\ w(k-1)\tilde{z}(k-1) \end{bmatrix} \\ &= w(k)\tilde{z}^2(k) + w(k-1)\tilde{z}^2(k-1) \end{aligned}$$

25. Solution to Least Squares Problem

Minimize objective function to find $\hat{\boldsymbol{\theta}}$:

$$\begin{aligned} \boldsymbol{J}[\hat{\boldsymbol{\theta}}(\boldsymbol{k})] &= \boldsymbol{w}(\boldsymbol{k})\tilde{\boldsymbol{z}}^2(\boldsymbol{k}) + \cdots + \boldsymbol{w}(\boldsymbol{k}-\boldsymbol{N})\tilde{\boldsymbol{z}}^2(\boldsymbol{k}-\boldsymbol{N}) \\ &= \tilde{\boldsymbol{Z}}(\boldsymbol{k})\boldsymbol{W}(\boldsymbol{k})\tilde{\boldsymbol{Z}}(\boldsymbol{k}) \\ &= [\boldsymbol{Z}(\boldsymbol{k}) - \hat{\boldsymbol{Z}}(\boldsymbol{k})]^T \boldsymbol{W}(\boldsymbol{k})[\boldsymbol{Z}(\boldsymbol{k}) - \hat{\boldsymbol{Z}}(\boldsymbol{k})] \end{aligned}$$

Minimize $\boldsymbol{J}$ and determine $\hat{\boldsymbol{\theta}}_{\mathrm{WLS}}$:

$$\hat{\boldsymbol{\theta}}_{\mathrm{WLS}}(\boldsymbol{k}) = \mathrm{arg}\,\min_{\boldsymbol{\theta}} \boldsymbol{J}[\hat{\boldsymbol{\theta}}(\boldsymbol{k})]$$

26. Solution to Least Squares Problem - Continued

Recall $\hat{\boldsymbol{\theta}}_{\text{WLS}}$ is obtained by minimizing

$$\boldsymbol{J}[\hat{\boldsymbol{\theta}}(\boldsymbol{k})] = [\boldsymbol{Z}(\boldsymbol{k}) - \hat{\boldsymbol{Z}}(\boldsymbol{k})]^T \boldsymbol{W}(\boldsymbol{k})[\boldsymbol{Z}(\boldsymbol{k}) - \hat{\boldsymbol{Z}}(\boldsymbol{k})]$$

$$\hat{\boldsymbol{Z}}(\boldsymbol{k}) = \boldsymbol{\Phi}(\boldsymbol{k})\hat{\boldsymbol{\theta}}(\boldsymbol{k})$$

Substituting for $\hat{\boldsymbol{Z}}(\boldsymbol{k})$,

$$\boldsymbol{J}[\hat{\boldsymbol{\theta}}(\boldsymbol{k})] = [\boldsymbol{Z}(\boldsymbol{k}) - \boldsymbol{\Phi}(\boldsymbol{k})\hat{\boldsymbol{\theta}}(\boldsymbol{k})]^T \boldsymbol{W}(\boldsymbol{k})[\boldsymbol{Z}(\boldsymbol{k}) - \boldsymbol{\Phi}(\boldsymbol{k})\hat{\boldsymbol{\theta}}(\boldsymbol{k})]$$

We drop the argument $\boldsymbol{k}$ temporarily for convenience and obtain,

$$\boldsymbol{J}[\hat{\boldsymbol{\theta}}] = \boldsymbol{Z}^T \boldsymbol{W} \boldsymbol{Z} - \boldsymbol{2}\boldsymbol{Z}^T \boldsymbol{W} \boldsymbol{\Phi} \hat{\boldsymbol{\theta}} + \hat{\boldsymbol{\theta}}^T \boldsymbol{\Phi}^T \boldsymbol{W} \boldsymbol{\Phi} \hat{\boldsymbol{\theta}}$$

To find $\hat{\boldsymbol{\theta}}$ at which $\boldsymbol{J}$ is minimum, differentiate and equate to zero:

$$\frac{\partial \boldsymbol{J}}{\partial \hat{\boldsymbol{\theta}}} = -\boldsymbol{2}\boldsymbol{\Phi}^T \boldsymbol{W} \boldsymbol{Z} + \boldsymbol{2}\boldsymbol{\Phi}^T \boldsymbol{W} \boldsymbol{\Phi} \hat{\boldsymbol{\theta}} = \boldsymbol{0}$$

27. Solution to Least Squares Problem - Continued

$$\frac{\partial J}{\partial \hat{\theta}} = -2\Phi^T W Z + 2\Phi^T W \Phi \hat{\theta} = 0$$

From this, we arrive at the normal equation,

$$\Phi^T W \Phi \hat{\theta} = \Phi^T W Z$$

Assume that $\Phi^T W \Phi$ is nonsingular. Persistence Condition.

$$\hat{\theta}_{\mathrm{WLS}}(k) = [\Phi^T(k) W(k) \Phi(k)]^{-1} \Phi^T(k) W(k) Z(k)$$